Calcolo (2016) 53:635–645
DOI 10.1007/s10092-015-0166-8

Iterative method for solving nonlinear singular problems

Agnieszka Prusińska[1] · Alexey Tret'yakov[2,3]

Received: 3 July 2015 / Accepted: 20 November 2015 / Published online: 10 December 2015

Abstract The aim of our work is to present a method (p-factor method) for solving nonlinear equations of the form

$$F(x) = 0, \quad F : \mathbb{R}^n \to \mathbb{R}^m,$$

in singular (irregular) case, i.e. when the matrix $F'(x)$ is singular at an initial point x_0 of an iterative sequence $\{x_k\}$, $k = 1, 2, \ldots$ We investigate conditions that have to be fulfilled at the initial point x_0 to obtain the existence of solution of nonlinear equation $F(x) = 0$, prove convergence of the presented p-factor method and give estimation of convergence rate for this method.

Keywords p-Regularity · Existence of solutions · p-Factor operator · Nonlinearity · Convergence

Research of the second author is supported by the Russian Foundation for Basic Research Grant No 14-07-00805 and the Council for the State Support of Leading Scientific Schools Grant 4640.2014.1 and by the Russian Academy of Sciences, Presidium Programme P-8.

✉ Agnieszka Prusińska
aprus@uph.edu.pl

Alexey Tret'yakov
tret@uph.edu.pl

1 Department of Mathematics and Physics, Siedlce University of Natural Sciences and Humanities, Konarskiego 2, 08-110 Siedlce, Poland

2 System Research Institute, Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland

3 Dorodnitsyn Computing Center, Russian Academy of Sciences, Vavilova 40, Moscow 119991, Russia

Mathematics Subject Classification 65H10 · 65J15 · 49N60 · 46T20

1 Introduction

We consider the problem of a construction of a numerical method (or an iterative sequence $\{x_k\}$, $k = 1, 2, \ldots$) for solving a system of nonlinear equations

$$F(x) = 0, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$ and F is singular at the initial point x_0. The problem (1) is called *regular* at the point x_0 if rank $F'(x_0) = m$. Otherwise, it is called *nonregular* (singular).

Usually, for proving convergence of the numerical method and obtaining an estimation of a convergence rate, a local convergence is supposed, i.e. the initial point x_0 of the iterative sequence is assumed to belong to some neighborhood $U_\varepsilon(x^*) = \{x : \|x - x^*\| < \varepsilon\}$, $\varepsilon > 0$ of a solution x^* to the system (1) and moreover this solution has to exist. For this reason, an important problem in construction numerical method one can faced is to prove the existence of solution in some neighborhood of the initial point and moreover to fix assumptions on a neighborhood of the starting point so that the constructed numerical method is convergent.

For the classical Newton (or Gauss–Newton) method in regular case,

$$x_{k+1} = x_k - \left(F'(x_0)\right)_R^{-1} F(x_k), \quad k = 1, 2, \ldots, \tag{2}$$

where $\left(F'(x_0)\right)_R^{-1} = F'(x_0)\left(F'(x_0)F'(x_0)^T\right)^{-1}$, the existence of a solution x^* in neighborhood of the initial point x_0 is guaranteed by Theorem 1 from the Sect. 2. And if the assumptions of Theorem 1 are valid, for the Newton method we have geometrical estimation of convergence rate:

$$\|x_{k+1} - x^*\| \leq q\|x_k - x^*\|, \quad k = 0, 1, 2, \ldots, \quad q \in (0, 1).$$

However, one of the main requirements for the method convergence, existence of solutions and estimation of the convergence rate is nonsingularity of the Jacobian matrix $F'(x_0)$. In a singular case, when $\left(F'(x_0)\right)_R^{-1}$ does not exist, the situation is much more complicated.

In the paper we present a numerical method (p-factor method) for solving equations in the case when $F'(x_0)$ is singular. We prove a theorem on existence of solutions and justify convergence and estimation of the rate of convergence to the described method.

One could wonder why we propose a new method instead of choosing another initial point and then using the classical method. It is because of the fact that even if one choose a starting point which is not singular it could happened that the next iteration might give rejection effect, that is $\|x_1 - x^*\| \gg \|x_0 - x^*\|$. The method described in this paper allows to avoid such effects and gives quadratic rate of convergence.

The construction of p-regularity introduced in [7] gives new possibilities for description and investigation of the set of solutions in the singular case. We use some of the methods to prove the conditions of the existence of local solutions to nonlinear singular equations in [4] and then some other results in the case with non-zero p-kernel (see [5]) and also give a description of a tangent cone in singular case [6].

Throughout this paper we assume that the mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ is continuously p-times differentiable on $\mathbb{R}^n$ and its pth order derivative at $x \in \mathbb{R}^n$ will be denoted as $F^{(p)}(x)$ (a symmetric multilinear map of p copies of $\mathbb{R}^n$ to $\mathbb{R}^m$) and the associated p-form is

$$F^{(p)}(x)[h]^p = F^{(p)}(x)[\underbrace{h, \ldots, h}_{p}].$$

Let X and Y are sets. We denote by 2^Y the set of all subsets of the set Y. Any mapping $\Phi : X \to 2^Y$ is said to be a *multimapping* (or a *multivalued mapping*) from X into Y. For a linear operator $\Lambda : X \to Y$, we define by Λ_R^{-1} its *right inverse*, that is $\Lambda_R^{-1} : Y \to 2^X$ and $\Lambda_R^{-1} y = \{x \in X : \Lambda x = y\}$. Of course, $\Lambda \Lambda_R^{-1} = I_Y$. Furthermore, we shall use the "norm" of such right inverse operator

$$\|\Lambda_R^{-1}\| = \max_{\|y\|=1} \min\{\|x\| : \Lambda x = y, \quad x \in X\}, \tag{3}$$

and so

$$\left\| \left(F^{(p)}(x_0) \right)_R^{-1} \right\| = \max_{\|y\|=1} \min \left\{ \|x\| : F^{(p)}(x_0)[x]^p = y, \quad x \in \mathbb{R}^n \right\}.$$

In our further considerations, under notion Λ^{-1} we shall mean just right inverse operator (multivalued) with the norm defined by (3).

The set $M(x^*) = \{x \in U : F(x) = F(x^*) = 0\}$ is called the *solution set* for the mapping F in neighborhood U of x^*.

We call $h \in \mathbb{R}^n$ *a tangent vector* to a set $M \subseteq \mathbb{R}^n$ at $x^* \in M$ if there exist $\varepsilon > 0$ and a function $r : [0, \varepsilon] \to \mathbb{R}^n$ such that $x^* + th + r(t) \in M$, where $t \in [0, \varepsilon]$ and $\|r(t)\| = o(t)$.

The collection of all tangent vectors at x^* is called *the tangent cone* to M at x^* and it is denoted by $T_1 M(x^*)$.

We recall two auxiliary lemmas. The first of these lemmas is a "multivalued" generalization of the contraction mapping principle. By $\mathrm{dist}(A_1, A_2)$ we mean the Hausdorff distance between sets A_1 and A_2.

Lemma 1 (Contraction multimapping principle [2]) *Assume that we are given a multimapping*

$$\Phi : U_\varepsilon(z_0) \to 2^{\mathbb{R}^n},$$

on a ball $U_\varepsilon(z_0) = \{z : \rho(z, z_0) \le \varepsilon\}$ $(\varepsilon > 0)$ *where the sets* $\Phi(z)$ *are non-empty and closed for any* $z \in U_\varepsilon(z_0)$. *Further, assume that there exists a number* θ, $0 < \theta < 1$ *such that*

1. $\operatorname{dist}(\Phi(z_1), \Phi(z_2)) \le \theta\rho(z_1, z_2)$ *for any* $z_1, z_2 \in U_\varepsilon(z_0)$
2. $\rho(z_0, \Phi(z_0)) < (1-\theta)\varepsilon$.

Then, for every number ε_1 *which satisfies the inequality*

$$\rho(z_0, \Phi(z_0)) < \varepsilon_1 < (1-\theta)\varepsilon,$$

there exists $z \in B_{\varepsilon_1/(1-\theta)}(z_0) = \{\omega : \rho(\omega, z_0) \le \varepsilon_1/(1-\theta)\}$ *such that*

$$z \in \Phi(z). \tag{4}$$

Lemma 2 [2] *Let* $\Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. *We set*

$$C(\Lambda) = \max_{\|y\|=1} \min\left\{\|x\| : x \in \mathbb{R}^n,\ \Lambda x = y\right\}.$$

If $\operatorname{Im}\Lambda = \mathbb{R}^m$, *then* $C(\Lambda) < \infty$.

2 Elements of *p*-regularity theory

Let us consider the case when the regularity condition fails but the mapping F is p-regular. We construct the p-factor operator under the assumption that $\mathbb{R}^m$ is decomposed into an orthogonal sum

$$\mathbb{R}^m = Y_1 \oplus \cdots \oplus Y_p, \tag{5}$$

where $Y_1 = \operatorname{Im} F'(x_0)$, and the remaining spaces are defined as follows. Let $Z_1 = \mathbb{R}^m$, $Z_2 = Y_1^\perp$, and let $P_{Z_2} : Y \to Z_2$ be the orthogonal projection onto Z_2. Let Y_2 be a linear span of the image of the quadratic map $P_{Z_2} F''(x_0)[\cdot]^2$. More generally, define inductively,

$$Y_i = \operatorname{span} \operatorname{Im} P_{Z_i} F^{(i)}(x_0)[\cdot]^i \subseteq Z_i, \quad i = 2, \ldots, p-1,$$

where $Z_i = (Y_1 \oplus \cdots \oplus Y_{i-1})^\perp$ with respect to $\mathbb{R}^m$, $i = 2, \ldots, p$ and $P_{Z_i} : \mathbb{R}^m \to Z_i$ is the orthogonal projection onto Z_i, $i = 2, \ldots, p$. Finally, $Y_p = Z_p$. The order p is the minimum number for which (5) holds.

Define the following mappings $f_i : U \to Y_i$, $f_i(x) = P_{Y_i} F(x)$, $i = 1, \ldots, p$, where $P_{Y_i} : \mathbb{R}^m \to Y_i$ is the orthogonal projection onto Y_i, $i = 1, \ldots, p$ (see e.g. [8]). It is obvious that $F(x) = f_1(x) + \cdots + f_p(x)$.

If $F^{(i)}(x_0) = 0$, where $i = 1, \ldots, p-1$ then we say that F is *completely degenerate* at $x_0 \in \mathbb{R}^n$ up to the order p. Note that $f_k^{(i)}(x_0) = 0$, $i = 1, \ldots, k-1$, $k = 1, \ldots, p$.

Definition 1 The linear operator $\Lambda_h \in \mathcal{L}(\mathbb{R}^n, Y_1 \oplus \cdots \oplus Y_p)$ is defined for $h \in \mathbb{R}^n$ by

$$\Lambda_h(x) = f_1'(x_0)x + f_2''(x_0)[h]x + \cdots + \frac{1}{(p-1)!} f_p^{(p)}(x_0)[h]^{p-1}x, \quad \text{for } x \in \mathbb{R}^n$$

and is called the *p-factor operator* along h at the point x_0.

Sometimes it is convenient to use the following equivalent definition of p-factor operator $\widetilde{\Lambda}_h \in \mathcal{L}\left(\mathbb{R}^n, Y_1 \times \cdots \times Y_p\right)$ for $h \in \mathbb{R}^n$,

$$\widetilde{\Lambda}_h(x) = \left(f_1'(x_0)x, f_2''(x_0)[h]x, \ldots, \frac{1}{(p-1)!} f_p^{(p)}(x_0)[h]^{p-1}x\right), \quad \text{for } x \in \mathbb{R}^n.$$

We also introduce the corresponding multivalued right inverse operator,

$$\Lambda_h^{-1} = \left\{x : y = \Lambda_h(x), y \in \mathbb{R}^m\right\},$$

and nonlinear operator

$$\Psi[h]^p = \left(f_1'(x_0)[h], f_2''(x_0)[h]^2, \ldots, \frac{1}{(p-1)!} f_p^{(p)}(x_0)[h]^p\right)$$

and $\Psi^{-1}(y) = \{h \in \mathbb{R}^n : y = \Psi[h]^p\}$. It is easy to see that $\Psi[h]^p = \widetilde{\Lambda}_h(h)$.

Let us introduce the following set

$$\mathrm{H}_p(x_0) := \left\{h : f_1'(x_0)h + f_2''(x_0)[h]^2 + \cdots + \frac{1}{(p-1)!} f_p^{(p)}(x_0)[h]^p = 0\right\}.$$

We can write it also as follows

$$\mathrm{H}_p(x_0) = \bigcap_{i=1}^{p} \mathrm{Ker}^i f_i^{(i)}(x_0),$$

where

$$\mathrm{Ker}^i f_i^{(i)}(x_0) = \left\{h \in \mathbb{R}^n : f_i^{(i)}(x_0)[h]^i = 0\right\}.$$

For our further considerations we need a generalization of the notion of regular mapping.

Definition 2 We say that the mapping F is *p-regular at x_0 along h* if $\mathrm{rank}\Lambda_h = m$.

Definition 3 We say that the mapping F is *p-regular at x_0* if either it is p-regular along every h from the set $\mathrm{H}_p(x_0) \setminus \{0\}$ or $\mathrm{H}_p(x_0) = \{0\}$.

3 Regular case

In regular case the tangent cone to a solution set of the Eq. (1) equals to the kernel of the first derivative of the mapping F, i.e. $T_1 M(x_0) = \mathrm{Ker} F'(x_0)$. To investigate regular nonlinear problems one can apply classical results, such as Lyusternik theorem, implicit function theorem, Lagrange–Euler optimality conditions.

Besides description of the solution set and formulation of optimality conditions, very important problem is to give a guarantee of existence of a solution in some neighborhood of an initial point.

We quote one of the modifications of the theorem about existence of solutions of Eq. (1) in regular case (see e.g. [1]).

Theorem 1 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$, $F \in \mathcal{C}^2(U_\varepsilon(x_0))$, $\varepsilon > 0$, $F'(x_0) \neq 0$, $\eta = \|F(x_0)\|$, F is regular at x_0 and $\delta = \left\|[F'(x_0)]^{-1}\right\|$, $c = \max_{x \in U_\varepsilon(x_0)} \|F''(x)\|$. If the following conditions*

1. $\delta \cdot c \cdot \varepsilon \leq \frac{1}{6}$,
2. $\delta \cdot \eta \leq \frac{\varepsilon}{2}$

are valid then the equation $F(x) = 0$ has a solution $x^* \in U_\varepsilon(x_0)$, the method (2) converges to x^* and the following estimation holds:

$$\|x_{k+1} - x^*\| \leq q\|x_k - x^*\|, \quad k = 1, 2, \ldots, q \in (0, 1).$$

For the proof see [4].

4 Singular case

We can generalize the above results and give the necessary conditions that have to be valid for existence of solutions in the singular case, i.e. when $F'(x_0)$ is singular. First, let us recall two lemmas that are essential in the proof of Theorems 2 and 3.

Lemma 3 [3] *Let U be an open subset of $\mathbb{R}^n$ such that $[a, b] \subset U$. If $f : U \to \mathbb{R}^m$ and $f \in \mathcal{C}^1(U)$ then*

$$\|f(b) - f(a) - \Lambda(b - a)\| \leq \max_{\xi \in [a,b]} \|f'(\xi) - \Lambda\| \cdot \|a - b\|, \text{ for any } \Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m).$$

Lemma 4 [4] *Let $S : \mathbb{R}^n \to \mathbb{R}^m = Y_1 \oplus \cdots \oplus Y_p$ be a nonlinear operator of the form $S(x) = \left(S_1(x), S_2(x), \ldots, S_p(x)\right)$, where $S_r(x)$ is r-form, $r = 1, \ldots, p$. If*

$$c = \max_{z \neq 0} \left\| S^{-1}\left(\frac{z}{\|z\|}\right)\right\| = \max_{z \neq 0} \min\left\{\|x\| : S(x) = \frac{z}{\|z\|}\right\} < \infty, \tag{6}$$

then for $y = \left(y_1, \ldots, y_p\right)$, $y_i \in Y_i$, $i = 1, \ldots, p$, where $\|y\| = \|y_1\| + \ldots + \|y_p\|$, we have

$$\left\|S^{-1}(y)\right\| \leq c \cdot p\left(\|y_1\| + \cdots + \|y_p\|^{\frac{1}{p}}\right). \tag{7}$$

For the singular case we have the following generalization of the Theorem 1.

Theorem 2 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$, $F \in \mathcal{C}^{p+1}(U_\varepsilon(x_0))$, $\varepsilon > 0$ and there exists h, such that $h \in \Psi^{-1}(-F(x_0)) \setminus \{0\}$, $\hat{h} = \frac{h}{\|h\|}$ and F is p-regular at x_0 along h. Moreover, assume that for a given ε the following conditions are satisfied*

1. $c_3 p^2 \delta^{\frac{1}{p}} \le \frac{1}{3}\varepsilon$,
2. $0 < \delta < 1$,
3. $\frac{4}{3}\left(4^p - 1\right) c_1 c_2 \varepsilon \le \frac{1}{2}$,

where $\delta = \|F(x_0)\| \ne 0$, $c_1 = \left\|\Lambda_{\widehat{h}}^{-1}\right\|$, $c_2 = \max_i \max_{x \in U_\varepsilon(x_0)} \left\| f_i^{(i+1)}(x)\right\| < \infty$, $i = 1, \dots, p$, $c_3 = \left\|\Psi^{-1}\right\| < \infty$. *Then the equation* $F(x) = 0$ *has a solution* $x^* \in U_\varepsilon(x_0)$.

For the proof see [4]. Notice that if $p = 1$ then Theorem 2 reduces to the Theorem 1. If $p \ge 2$ then $F'(x_0)$ is singular and we have singular case.

Remark 1 (Sequential p-regularity) The assumptions of Theorem 2 can be relax if we consider the sequence:

$$\xi_0 = 0, \xi_{k+1} = \xi_k - \Lambda_h^{-1}\left(f_1(x_0 + h + \xi_k) + \dots + f_p(x_0 + h + \xi_k)\right), \quad k = 1, 2, \dots.$$

Namely, instead of p-regularity of F at x_0 we can assume p-regularity F on $\{\xi_k\}$, $k = 1, 2, \dots$. It means that instead of surjectivity of Λ_h we can require that $\mathrm{Im}\Lambda_h = Y_1' \oplus \dots \oplus Y_p'$ and $f_i(x_0 + h + \xi_k) \in Y_i'$, where $Y_i' \subseteq Y_i$, $i = 1, \dots, p$, $k = 0, 1, \dots$.

Remark 2 We choose $x \in \mathbb{R}^n$ as a representative of the set $\Lambda_h^{-1}(y)$ if $\|x\| = \min\{\|z\| : \Lambda_h(z) = y\}$. We will write then $\Lambda_h^{-1}(y) = x$.

Consider the following method (p-factor method)

$$z_0 = x_0 + h, \; z_{k+1} = z_k - \Lambda_h^{-1}(F(z_k)), \quad k = 0, 1, 2, \dots. \tag{8}$$

Theorem 3 *Suppose that for* F, x_0, h, ε, δ *the assumptions of Theorem 2 are valid. Then the sequence defined in (8) converges to the solution* $x^* \in U_\varepsilon(x_0)$ *and the following estimation of convergent rate holds:*

$$\|z_{k+1} - x^*\| \le q\|z_k - x^*\|, \quad k = 0, 1, 2, \dots, q \in (0, 1). \tag{9}$$

Proof Consider the following equivalent form of the sequence (8).

$$\xi_0 = 0, \; \xi_{k+1} = \xi_k - \Lambda_h^{-1}(F(x_0 + h + \xi_k)), \quad k = 0, 1, 2, \dots.$$

We prove that $\forall_k \|\xi_{k+1} - \xi_k\| \le \frac{1}{2}\|\xi_k - \xi_{k-1}\|$. For this purpose we define a multivalued mapping $\Phi_h : U_\varepsilon(x_0) \to 2^{\mathbb{R}^n}$, such that

$$\begin{aligned} \Phi_h(x) &\in x - \Lambda_h^{-1}(f_1(x_0 + h + x) + \dots + f_p(x_0 + h + x)) \\ &= x - \Lambda_h^{-1}(F(x_0 + h + x)), \quad x \in U_\varepsilon(x_0). \end{aligned}$$

The sets $\Phi_h(x)$ are non-empty because Λ_h is a surjection for any $x \in U_\varepsilon(x_0)$. Moreover, for any $y \in Y_1 \times \dots \times Y_p$ the sets $\Lambda_h^{-1}(y)$ are linear manifolds parallel to $\mathrm{Ker}\Lambda_h$, and hence the sets $\Phi_h(x)$ are closed for any $x \in U_\varepsilon(x_0)$. Taking into account Remark 2 we can write $\|\xi_{k+1} - \xi_k\| = \operatorname{dist}(\Phi_h(\xi_{k+1}), \Phi_h(\xi_k))$. We prove that

$$\operatorname{dist}(\Phi_h(\xi_{k+1}), \Phi_h(\xi_k)) \leq \frac{1}{2}\|\xi_k - \xi_{k-1}\|, \tag{10}$$

for $\xi_{k+1}, \xi_k \in U_{\frac{\varepsilon}{2}}(x_0)$ such that $\|\xi_j\| \leq \frac{\|h\|}{R}$, $j = k, k+1$, where

$$R = \max_i R_i,\ R_i = \max\left\{1, \frac{2}{\varepsilon \cdot c_2} \cdot \frac{1}{(i-1)!}\left\| f_i^{(i)}(x_0)\right\|\right\}, \quad i = 1, \ldots, p.$$

Let $\Lambda_{h,i} = \frac{1}{(i-1)!} f_i^{(i)}(x_0)[h]^{i-1}$, $i = 1, \ldots, p, s_1 = x_0 + h + \xi_k, s_2 = x_0 + h + \xi_{k-1}$. Then

$$\begin{aligned}
&\operatorname{dist}(\Phi_h(\xi_{k+1}), \Phi_h(\xi_k)) = \inf\{\|z_1 - z_2\| : z_1 \in \Phi_h(\xi_k),\ z_2 \in \Phi_h(\xi_{k+1})\}\\
&\quad = \inf\left\{\|z_1 - z_2\| : \Lambda_h(z_{i+1}) = \Lambda_h(\xi_{k-i}) - \left(f_1(s_{i+1}) + \cdots + f_p(s_{i+1})\right),\ i = 0, 1\right\}\\
&\quad \leq \inf\left\{\|z\| : \Lambda_h(z) = \Lambda_h(\xi_k - \xi_{k-1}) - \left(f_1(s_1) - f_1(s_2) + \cdots + f_p(s_1) - f_p(s_2)\right)\right\}\\
&\quad = \inf\Bigg\{\|z\| : \Lambda_h(z) = \Bigg(\Lambda_{h,1}(\xi_k - \xi_{k-1}) - f_1(s_1) + f_1(s_2)\\
&\qquad + \cdots + \frac{1}{\|h\|^{p-1}}\left(\Lambda_{h,p}(\xi_k - \xi_{k-1}) - f_p(s_1) + f_p(s_2)\right)\Bigg)\Bigg\}\\
&\quad \leq \inf\Bigg\{\|z\| : z = \Lambda_h^{-1}\Bigg(\Lambda_{h,1}(\xi_k - \xi_{k-1}) - f_1(s_1) + f_1(s_2)\\
&\qquad + \cdots + \frac{1}{\|h\|^{p-1}}\left(\Lambda_{h,p}(\xi_k - \xi_{k-1}) - f_p(s_1) + f_p(s_2)\right)\Bigg)\Bigg\}\\
&\quad \leq c_1 \cdot \sum_{i=1}^{p} \frac{1}{\|h\|^{i-1}}\left\| f_i(s_1) - f_i(s_2) - \Lambda_{h,i}(s_1 - s_2)\right\|.
\end{aligned}$$

Taking into account Lemma 3, we have

$$\begin{aligned}
&\left\| f_i(s_1) - f_i(s_2) - \Lambda_{h,i}(s_1 - s_2)\right\|\\
&\quad \leq \sup_{\theta \in [0,1]}\left\| f_i'(s_2 + \theta(s_1 - s_2)) - \Lambda_{h,i}\right\| \cdot \|\xi_k - \xi_{k-1}\|.
\end{aligned} \tag{11}$$

As f_i is completely degenerate up to the order i we obtain the following Taylor expansion

$$\begin{aligned}
&f_i'(s_2 + \theta(s_1 - s_2))\\
&\quad = f_i'(x_0) + \ldots + \frac{f_i^{(i)}(x_0)[s_2 - x_0 + \theta(s_1 - s_2)]^{i-1}}{(i-1)!} + \omega_i(h, \xi_k, \xi_{k-1}, \theta)\\
&\quad = \frac{f_i^{(i)}(x_0)[s_2 - x_0 + \theta(s_1 - s_2)]^{i-1}}{(i-1)!} + \omega_i(h, \xi_k, \xi_{k-1}, \theta),
\end{aligned} \tag{12}$$

where

$$\|\omega_i(h, \xi_k, \xi_{k-1}, \theta)\| \leq \sup_{x \in U_\varepsilon(x_0)} \left\| f_i^{(i+1)}(x)[h + \xi_{k-1} + \theta(s_1 - s_2)]^i \right\|.$$

Moreover,

$$\begin{aligned} f_i^{(i)}(x_0)[s_2 - x_0 + \theta(s_1 - s_2)]^{i-1} &= f_i^{(i)}(x_0)[h + \xi_{k-1} + \theta(s_1 - s_2)]^{i-1} \\ &= \sum_{t=0}^{i-1} \binom{i-1}{t} f_i^{(i)}(x_0)[h]^{i-1-t}[\xi_{k-1} + \theta(s_1 - s_2)]^t \\ &= f_i^{(i)}(x_0)[h]^{i-1} + \sum_{t=1}^{i-1} \binom{i-1}{t} f_i^{(i)}(x_0)[h]^{i-1-t}[\xi_{k-1} + \theta(s_1 - s_2)]^t. \end{aligned} \tag{13}$$

Thus, inserting (14) into (13) and then putting the obtained formula into (11), we get

$$\begin{aligned} &\left\| f_i(s_1) - f_i(s_2) - \Lambda_{h,i}(s_1 - s_2) \right\| \\ &\leq \sup_{\theta \in [0,1]} \left\| \frac{\sum_{t=1}^{i-1} \binom{i-1}{t} f_i^{(i)}(x_0)[h]^{i-1-t}[\xi_k + \theta(s_1 - s_2)]^t}{(i-1)!} + \omega_i(h, \xi_k, \xi_{k-1}, \theta) \right\| \\ &\cdot \|\xi_k - \xi_{k-1}\|. \end{aligned} \tag{14}$$

Let $F(x_0) = (y_1, \ldots, y_p)$, where $y_i \in Y_i,\ i = 1, \ldots, p$. Then from the assumption and the definition of norm in $\mathbb{R}^n$ we have $\|y_1\| + \cdots + \|y_p\| \leq \delta$. As $R\|\xi_j\| \leq \|h\|$, $j = k-1, k$ we have $\|h + \xi_{k-1} + \theta(s_1 - s_2)\| \leq 4\|h\|$. Taking into account Lemma 4 and assumptions (3) and (2) of Theorem 2 we have

$$\begin{aligned} \|h\| &\leq (1 + \Delta)\|\Psi^{-1}(-F(x_0))\| \leq (1 + \Delta)c_3 p \left(\|y_1\| + \cdots + \|y_p\|^{\frac{1}{p}}\right) \\ &\leq (1 + \Delta)c_3 p^2 \delta^{\frac{1}{p}} \leq \frac{\varepsilon}{2}, \end{aligned}$$

where $0 < \Delta < \frac{1}{2}$. This and the previous formulae imply

$$\|\omega_i(h, \xi_k, \xi_{k-1}, \theta)\| \leq c_2\|h + \xi_{k-1} + \theta(s_1 - s_2)\|^i \leq 4^i c_2 \frac{\varepsilon}{2} \|h\|^{i-1}. \tag{15}$$

Moreover,

$$\|\xi_{k-1} + \theta(s_1 - s_2)\| \leq 3\|h\|/R \leq 3\|h\|/R_i. \tag{16}$$

Taking into account the definition of R_i, we get

$$\begin{aligned}
&\left\| \sum_{k=1}^{i-1} \binom{i-1}{k} f_i^{(i)}(x_0)[h]^{i-1-k}[\xi_{k-1} + \theta(s_1 - s_2)]^k \right\| \\
&\quad \le \left\| f_i^{(i)}(x_0) \right\| \cdot \sum_{k=1}^{i-1} \binom{i-1}{k} \|h\|^{i-1-k} (3\|h\|)^k / R_i^k \\
&\quad \le \left\| f_i^{(i)}(x_0) \right\| \cdot \|h\|^{i-1} \cdot 4^{i-1} / R_i \le 4^i (i-1)! \frac{\varepsilon}{2} c_2 \|h\|^{i-1}.
\end{aligned} \tag{17}$$

Now, inserting (15)–(17) into (14) we obtain

$$\left\| f_i(s_1) - f_i(s_2) - \Lambda_{h,i}(s_1 - s_2) \right\| \le 4^i \varepsilon c_2 \|h\|^{i-1} \cdot \|\xi_k - \xi_{k-1}\|.$$

Hence

$$\begin{aligned}
\|\xi_{k+1} - \xi_k\| &\le c_1 \cdot \sum_{i=1}^{p} \frac{1}{\|h\|^{i-1}} 4^i c_2 \varepsilon \|h\|^{i-1} \|\xi_k - \xi_{k-1}\| \\
&= \frac{4}{3} \left(4^p - 1\right) c_1 c_2 \varepsilon \|\xi_k - \xi_{k-1}\| \le \frac{1}{2} \|\xi_k - \xi_{k-1}\|
\end{aligned}$$

which proves (10).

Let ξ_1 be an element of $\Phi_h(\xi_0)$ such that $\xi_1 \in \xi_0 - \Psi^{-1}(F(x_0))$ and $\|\xi_1 - \xi_0\| \le (1 + \Delta)\|\Psi^{-1}(-F(x_0)\|$, where $0 < \Delta < \frac{1}{2}$. Thus we have

$$\|\Phi_h(\xi_0) - \xi_0) \le \|\xi_1 - \xi_0\| \le (1 + \Delta) c_3 p^2 \delta^{\frac{1}{p}} \le \frac{1}{2}\varepsilon.$$

From the above and from (10) we obtain that the sequence ξ_k is a Cauchy sequence hence there exists an element ξ^* such that $\xi^* \in \Phi_h(\xi^*)$. It means that

$$0 \in \Lambda_h^{-1} \left(f_1(x_0 + h + \xi^*) + \cdots + f_p(x_0 + h + \xi^*) \right)$$

and hence $\left(f_1(x_0 + h + \xi^*) + \cdots + f_p(x_0 + h + \xi^*) \right) = \mathbf{0}$. Then for $i = 1, \ldots, p$, $f_i(x_0 + h + \xi^*) = 0$ which is equivalent to $F(x_0 + h + \xi^*) = \mathbf{0}$. It follows that the sequence $z_k = x_0 + h + \xi_k$ is convergent to $x^* = x_0 + h + \xi^*$ which is a solution of $F(x) = 0$.

The estimation (9) follows from (10). □

We conclude with an example which is very easy but serves to illustrate the main idea of the method.

Example

Consider Eq. (1) and let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be the mapping given by

$$F(x^1, x^2) = \begin{pmatrix} x^1 + x^2 + (x^1+1)^3 + 2 - 10^{-3} \\ x^1 x^2 + x^1 + x^2 + (x^2+1)^3 + 1 + 10^{-6} \end{pmatrix}.$$

Let $x_0 = (-1, -1)^T$ be an initial point. One can easily verify that F is nonregular at x_0. Indeed, $F'(x^1, x^2) = \begin{pmatrix} \frac{1+3(x^1+1)^2}{x^2+1} & \frac{1}{x^1+3(x^2+1)^2+1} \end{pmatrix}$ and $\operatorname{rank} F'(-1, -1) \neq 2$, so the Newton method could not be applied. We will prove that F is 2-regular and construct 2-factor operator.

We can decompose the space $\mathbb{R}^2$ as follows $\mathbb{R}^2 = Y_1 \oplus Y_2$, where $Y_1 = \left\{\binom{t}{0} : t \in \mathbb{R}\right\}$ and $Y_2 = \left\{\binom{0}{r} : r \in \mathbb{R}\right\}$. Then we calculate norms of some elements to examine if the assumptions of Theorem 2 are satisfied. We obtain $\delta = \frac{\sqrt{10^6+1}}{10^6}$, $\hat{h} = \left(\frac{1-\sqrt{3}}{2\sqrt{2}} \; \frac{1+\sqrt{3}}{2\sqrt{2}}\right)^T$ and 2-factor operator is $\Lambda_{\hat{h}} = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{3}}{2\sqrt{2}} & \frac{1-\sqrt{3}}{2\sqrt{2}} \end{pmatrix}$. Hence, $\operatorname{rank} \Lambda_{\hat{h}} = 2$, so it is surjective. To verify all assumptions of Theorem 2 we calculate $c_1 \approx 1.414214$, $c_2 \approx 1$, $c_3 \approx 0.001$, $\varepsilon = 0.01$ and hence we can conclude that in ε-neighbourhood of x_0 there exists a solution of (1).

We use the sequence (8) to find an approximation of the solution. Here $\xi_1^1 = 0$, $\xi_1^2 = 0$ and one of the three solutions which is in ε-neighbourhood of x_0 is $x^* \approx (-1.000619933, -99838)$.

In the table one can see the six iterations of the method described in this paper.

k	z_k^1	z_k^2	$z_k^1 - x^{1*}$	$z_k^2 - x^{2*}$
1	0	0	0.000619933	−0.001619933
2	−1.000366025	−0.998634	0.000253908	−0.00025393
3	−1.000366528	−0.998344	0.000253405	3.6067×10^{-5}
4	−1.00065657	−0.998343	-3.6637×10^{-5}	3.73067×10^{-5}
5	−1.000657155	−0.998392	-3.7222×10^{-5}	-1.1933×10^{-5}
6	−1.000608426	−0.998384	1.1507×10^{-5}	-3.933×10^{-6}

References

1. Demidovitch, B.P., Maron, I.A.: Basis of Computational Mathematics. Nauka, Moscow (1973)
2. Ioffe, A.D., Tihomirov, V.M.: Theory of Extremal Problems. Nauka, Moscow (1974). (in Russian)
3. Krasnosel'skii, M.A., Wainikko, G.M., Zabreiko, P.P., Rutitcki, YuB, Stetsenko, VYu.: Approximated Solutions of Operator Equations, p. 484. Walters-Noordhoff Publishing, Groningen (1972)
4. Prusińska, A., Tret'yakov, A.: A remak on the existence of solutions to nonlinear equations with degenerate mappings. Set-Valued Anal. **16**, 93–104 (2008)
5. Prusińska, A., Tret'yakov, A.: On the existence of solutions to nonlinear equations involving singular mappings with non-zero p-kernel. Set-Valued Anal. **19**, 399–416 (2011)
6. Prusińska, A., Tret'yakov, A.: *p*-Regularity theory. Tangent cone description in singular case. Ukr. Math. J. **67**,(8), 1097–1106 (2015)
7. Tret'yakov, A.: Necessary and sufficient conditions for optimality of p-th order. USSR Comput. Math. Math. Phys. **24**(1), 123–127 (1984)
8. Tret'yakov, A., Marsden, J.E.: Factor-analysis of nonlinear mappings: p-regularity theory. Commun. Pure Appl. Math. **2**, 425–445 (2003)